

DOCUMENT RESUME

ED 216 033

TM 820 189

AUTHOR Johanningmeier, E. V.
 TITLE Stuart Appleton Courtis: Tester, Reformer and Progressive.
 PUB DATE 22 Mar 82
 NOTE 55p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC03 Plus Postage.
 DESCRIPTORS *Biographies; *Educational Researchers; *Educational Testing; Elementary Secondary Education; Measurement Techniques; *Progressive Education; *Standardized Tests; *Test Theory
 IDENTIFIERS *Courtis (Stuart Appleton)

ABSTRACT The career of Stuart Appleton Courtis in the growth of testing and educational measurement parallels the development of progressive education in the first half of the twentieth century. In 1909 he developed the standardized Courtis Arithmetic Test, Series A, the first objective test used in any city public schools. Continuing his work in testing, he revised some of his earlier concepts in constructing tests for areas including reading and composition. His effort was to improve public education by challenging traditional conceptions with new practices based on the social and behavioral sciences. Discussion of his career covers his service as a director in educational research and the evolution of the tests he developed. The testing theories and formula Courtis used are presented. Participating in several school surveys, and writing texts in educational philosophy, which he taught at the University of Michigan, he saw himself as a progressive educator attempting to develop techniques for educational cooperation which he believed would improve the democratic process. He lived from 1874-1969.
 (CM)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED216033

Stuart Appleton Courtis:
Tester, Reformer and Progressive

by

E. V. Johanningmeier
University of South Florida
Tampa, Florida

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

E. V. Johanningmeier

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Presented at Annual Meeting of the American Educational
Research Association, Division F. Session 40.06: Educational
Leaders of the Progressive Era, March 22, 1982, New York.

Stuart Appleton Courtis:
Tester, Reformer and Progressive

by

E. V. Johanningmier
University of South Florida
Tampa, Florida

Were this a meeting of the Modern Language Association, a claim could be made that Stuart Appleton Courtis (1874-1969) was a minor poet or perhaps even a major poet who now deserves to have his place in literature defined. But Courtis was no poet. He was an educator and an early developer of standardized tests who has received little or no attention in our histories. Yet, to study his work is to study that overstudied but still engaging topic--progressive education--as well as the curious but usually ignored link between progressive education and the measurement movement in American education.

As a director of a bureau of educational research, as a test developer, as a student of measurement techniques, as a member of several school surveys, and as a teacher, Courtis participated in the twentieth century transformation of the

public school. Like others of his era, he eventually directed his interests and efforts not only to the development of ways to measure what schools and students accomplished but also to the determination of their social and political purposes. He saw himself as a progressive, tried to develop a technique for cooperation that he believed would improve the democratic process, and wrote philosophy of education texts for the courses he taught at the University of Michigan. He is among those who tried to build a "new education" in the first half of the twentieth century. He is a progressive because he is among those who tried to improve public education by challenging traditional conceptions and practices with new ones fashioned from science.

Thus, in this paper progressive education is viewed as the attempt to discard the traditional foundation of educational practices and replace it with one from the social and behavioral sciences. The focus of this definition is on the reasons given for an educational practice, not the practice itself. Progressive education was not a creed and it was not an event; it was a development. As educators turned to the sciences, especially psychology and sociology, either to defend earlier changes in educational practice or to improve the efficiency of public education, there developed a fundamental shift in the perspective of many educators.

Had Courtis taught for a year of two and then gone off to study with one of the major psychologists of the time, comparisons with Bagley and others could be continued. But Courtis began teaching even before he had earned his B.S. at Teachers College, Columbia University in 1919. His academic degrees (M.A. at Teachers College in 1921 and a Ph.D. at the University of Michigan in 1925) were earned several years after he had set the course for his career and made one of his major contributions to education.

By 1906 Courtis had been appointed head of the science and mathematics department at the Liggett School. Soon thereafter C. W. Stone published his arithmetic tests, and Courtis administered them to the students under his charge. As Arthur B. Moehalman told the story, "they revealed, as tests always do, an enormous range of individual differences in every grade."³ Courtis examined the results, concluded that he was not a very good teacher and offered his resignation. His principal, Emma Liggett, refused it and instructed him to find out why the differences existed. He then studied the works of Judd and Thorndike and developed his own tests which were published in 1909.

The Courtis Arithmetic Tests, Series A were used in the 1911 New York City Survey and thus became the first objective tests to be used in any city public school survey. Thirty thousand students then took his test. Administration

of tests to assess the quality, or efficiency, of teaching constituted a new form of evaluation. It was, perhaps, consistent with the tradition of superintendents who evaluated their teachers by examining a set of students' papers but was certainly very different from the practice of offering evaluations after an observation or two. Leonard P. Ayres later observed that the use of the Courtis tests "which had by that time attracted a good deal of attention" not only "settled all doubts as to the availability of the tests themselves for the measurement of educational attainment" but also "established the principle that in conducting school surveys scientific tests must be utilized where they are available."⁴

In 1916 Courtis reported that between August, 1914 and August, 1915, 455,007 of his tests had been sent to educators in 42 states.⁵ He soon constructed tests for several other areas: Composition Tests (1916); Series R, Reading Tests (1917); Series W, Writing Tests (1918); Series S, Spelling Tests (1918); Test A, Geography, Supervisory with B. A. Barnes (1918); Standard Samples of Handwriting (1919); Standard Samples of Compositions (1919); Test B, Geography, Supervisory with B. A. Barnes (1919); Courtis Standard Practice Tests in Arithmetic published by the World Book Company (1920); Courtis Standard Practice Tests in Handwriting with Lena M. Shaw

also published by the World Book Company (1922); Teacher Rating (1922); Series M, Music Tests (1922) and the Courtis Standard Research Tests (1920-34).⁶

Courtis' arithmetic tests (Series A and Series B) were found to be unsatisfactory for the Cleveland Survey, but, as George S. Counts reported in his dissertation, Courtis' help was secured in the construction of an appropriate test. There was, it seems, a gap between Series A and Series B--a gap between "the very simple type of example in the first series and the highly complex type in the second series." Courtis helped to develop "tests representing types of intermediate complexity"--a development that Counts described as "the logical evolution of Mr. Courtis' own system." "Whatever merit" the new test had, was, calimed Counts, due to Courtis' contributions.⁷ Counts used the data from the administration of the tests in Cleveland and Grand Rapids to complete his dissertation and to conclude with proper qualifications that

. . . the American children are not prodigies when compared with the children of other races. With the exception of the records for Grade 7-I the Hollanders and Americans are in close agreement, and as compared with the other races they

seem to be somewhat inferior. The Germans and the Swedes are intermediate groups, while the Slavs really seem to superior.⁸

Counts' study did not attempt to explain whether the superiority of Slaves in arithmetic was due to "biological or social factors."

When Walter S. Monroe undertook to give an account of then existing "tests and standards" in 1918, he included several of Courtis' tests. Courtis' Arithmetic Tests, Series A, he reported, had been replaced by Series B in 1914. Series A was an eight part test that included one test for each of the four fundamental operations for the number combinations 0 to 9 as well as tests on copying figures, speed reasoning, fundamental operations, and reasoning. Series B was "used extensively" and "proved reliable in 75 to 90 percent of the cases." Abbreviated forms of the Series B tests were used in the National Business Ability Tests. However, according to Monroe, the Series B tests did not allow as detailed an analysis as did the Cleveland Survey Tests developed by Judd and did not offer "carefully graded" examples "arranged in order of difficulty."⁹ Monroe had apparently overlooked the comments of Judd's student, Counts.

Courtis' Handwriting, Series W Test was "an untimed 'maximum performance' test, designed to secure samples of the children's best writing after practice." It employed the Ayres scale to measure the quality of the work, and used material that afforded "an analysis of the defects in the writing of a particular child."¹⁰ His Standard Research Tests in English had been designed to measure both speed and comprehension in silent reading but was discontinued because it "was so complex that the marking of the test papers was a laborious task."¹¹ His Silent Reading Test measured reading ability in Grades II, III, and IV and afforded measures for rate and comprehension. Students only had to give answers by marking "yes" or "no" so scoring was easy. The two forms of it allowed teachers to determine easily how much progress had been made by students between the beginning and the end of the school term.¹²

Courtis' Spelling Test used words from the Ayres Scale that were placed in sentences, "and the sentences are arranged so that they can be dictated at specified rates, which correspond to the rate of writing in several grades." The standards for this test were "slightly lower" than those set for Ayres' Standards.¹³

According to Moehlman, Courtis sold more than thirteen million copies of his tests by 1925. However, after

completing his dissertation--"Why Children Succeed"--he concluded "that the mathematical assumptions underlying educational measurement, including his own, were wrong." He then "scrapped his tests and continued more extensive research in Detroit schools and in certain foreign countries."¹⁴ In 1946 Courtis reported that "between 1909, when my first tests were published, and 1938, when I withdrew all my tests from the market because I had discovered that they did not measure what they were supposed to measure, I sold more than twenty million copies around the world."¹⁵ He had concluded that tests did not measure ability but only performance at a specified time.

Courtis' early work on standardized tests places him in the company of Stone, Thorndike, Ayres, Starch and many others who were developing instruments to measure objectively school and student progress. While Courtis and others were developing their measurement instruments, there was widespread belief in the utility of such efforts. Measurement was becoming an essential ingredient in the science of education. As Courtis wrote in 1916, it was an aspect of science and was therefore a worthy pursuit. It was "a method of proved worth in education." Those who declared that measurement was a "failure" were confessing their "own lack of confidence" or their "own incompetency."¹⁶

In 1914 Courtis left the Liggett School to become the Director of Educational Research in the Detroit Public Schools (1914-1919). In that post he "was turned loose among two thousand teachers and seventy thousand students to attempt to improve the efficiency of teaching of the fundamental subjects."¹⁷ From then on Courtis work was nearly evenly divided between work in school systems and higher education. From 1920 to 1924 he served as Director of Instruction and Dean at the Detroit Teachers College and as an educational consultant to the Detroit Public Schools (1924 to 1931) and to the Hamtramck Public Schools (1926 to 1930). From 1931 to 1944 he taught at Wayne University and from 1924 to 1944 at the University of Michigan where he was appointed Professor Emeritus upon his retirement.

Although Courtis' positions were confined to Michigan, his work and influence extended beyond his home state. Besides serving on the Hanus Committee on School Inquiry, New York City, 1911, he participated in the resurvey of the New York City Public Schools in 1924 and in the Gary, Indiana Survey in 1916, and was in charge of testing in the Boston Public Schools in 1916. He was president of the American Association for the Advancement of Science, 1918-1919 as well as a contributing member of the National Society for the Study of Education and the National Society of College Teachers of Education.

Courtis's first significant work--the 1909 publication of his arithmetic tests, his positions, and his publications all show that he was devoted to the development and practice of educational measurement. Through educational measurement education could, he and others believed, be improved significantly. During the course of his career, he changed some of his notions about educational measurement, became somewhat critical of the movement and its practices, but never abandoned his belief in the ultimate utility of measurement. His claims about the advantages and limitations of educational measurement and his views on educational research deserve attention. That Courtis did identify with the measurement movement and that he did accept the progressive movement may serve to remind us that the two are separate only in our studies. In fact, they are related not separate movements. Those who measured were not necessarily opposed to progressive education, although some did change their views or move away from their early interests in measurement. Harold O. Rugg is an obvious example; George S. Counts is certainly another.¹⁸

How we have managed to assume, usually tacitly, that progressive education and the measurement movement are separate, antithetical, or contrary movements remains an enigma. Perhaps the explanation is to be found in the

early 1960 s when we revised our approaches to the study of the history of education. Perhaps the explanation can be found in what I term the Callahan-Cremin legacy. In 1962 each of these historians of education published works that we still use as starting points for many of our inquiries. In The Transformation of the School Lawrence A. Cremin explained to the critics of American education during the post World War II era why the schools were as the critics claimed. He showed how educational theory had become divorced from the academic disciplines. His account was widely accepted and had considerable influence on how we conceptualized and acted upon problems attendant to public schooling. In Education and the Cult of Efficiency Raymond E. Callahan also undertook an exploration of the charges of the post World War II critics of American education. His focus was not, however, on the theoreticians of education but on the public school administrators. He related how the school administrators embraced Taylorism and applied principles of "scientific management" to the administration and organization of public schools. As they did so, they overlooked the traditional academic values and practices. Callahan's suggestion that school administrators be well trained in the social sciences was not inconsistent with Cremin's suggestion that educators needed to move close to their

"parent disciplines." Each was looking at the same problem and each was giving an account of the same period. However, each looked at different parts of the system. Neither should be criticized for what they did. However, it is curious that we have typically not observed how those two works fit together and consider the same period and set of problems. As we continued to explore the society that supported and influenced public education and turned to the "parent disciplines" for clues and methods, we simply ceased to look at our own developments in professional education.

A study of Courtis' views is also a study of the beginnings of that area of our professional field called educational research. Courtis is a good starting point for such an exploration. As a member of the National Association of Directors of Educational Research, he knew Joseph Mayer Rice and was one of the contributors to the first issue of the Journal of Educational Research (1920). At the end of his career he offered assessments of his work and the field with which he identified. They are sometimes reflective and critical. For example, in an article written for the 1938 NSSE Yearbook he noted how educational researchers had been influenced by the interests of engineers and upon how crude the first educational measures were:

The revelations of these first crude comparative testings, as with all subsequent work of this character, were interpreted as evidence of the inefficiency of school teaching. For the range of score in each type of test and in each grade was, and is, enormous, and the overlappings from grade to grade vastly exceed the average annual progress from grade to grade. At the time efficiency was an ideal widely accepted in the engineering world; for a period it also dominated the thinking of investigators in education. One school survey after another brought new 'deficiencies' to light and emphasized the range of individual differences in achievements at all levels and subjects.¹⁹

In 1942 Courtis even cited his own work on the New York Survey as an example of "crude comparative testings" that were used to show how inefficient teaching was. He was even more explicit about the inadequacy of his own work. Then he offered a quotation from the New York Survey Report that he had written. He had constructed a table to display the range of scores in arithmetic and then wrote that the table

. . . points out plainly facts of the largest significance both for education generally and for the local situation. It is the distributions of the individual scores that are startlingly--yes, sensationally significant. The real meaning is that, so far as any individual child is concerned, to say that he has completed the course in arithmetic in the public schools is to convey no information as to his ability in even the simplest work. He may be almost an absolute incompetent so far as practical work is concerned, or he may have acquired a degree of skill that would be adequate for any situation in which he is likely to find himself. As a whole, therefore, the table more than justifies the criticism that has yet been made by the 'man on the street.'²⁰

Over thirty years after having written that statement, Courtis described its content as "silly" and confessed that "it is humiliating to have written such pronouncements."¹⁹

Courtis' view of the history of educational measurement and its distinctive phases is also of interest, for there is, as a review of the literature demonstrates, no agreement on how to periodize the history of educational research.

For example, Walter S. Monroe chose 1918 as the beginning of the modern period in educational research and dismissed earlier work as simply belonging to the pioneer period. His position is illustrative of a longstanding, but false, belief that intelligence tests are an essential ingredient in educational research. According to Monroe, "the fact that there were only very limited formal provisions for educational research until near the close of this period, that practically no achievement tests were available until 1915, and that group intelligence tests were not available until after 1918 probably suggests that relatively little significant research was completed before 1918."²¹

Monroe's position reflects a bias against those who completed important studies that had lasting influence on education but were completed before 1918. Rice's work, conducted in the 1890 s, is one obvious example. While Rice did not have command of sophisticated statistical procedures, he did develop methods to gather "objective" data and did consider an array of variables that have been consistently considered by educational researchers since he presented his conclusions.²² Other examples include the Thorndike-Woodworth transfer studies (1901) and the series of studies conducted by Starch and E. C. Elliott on the reliability of grading in high school English, history and mathematics that appeared in the School Review (1912-13).

Monroe's position also overlooks the development of the questionnaire method by G. Stanley Hall and its origins and use in German pediatric research in the 1870 s.²⁴ It also ignores the influence of Sir Francis Galton and Franz Boas in the development of psychometric and statistical techniques. More importantly, Monroe's position overlooks the development and widespread use of several achievement tests--tests that were widely used in school surveys five years before the development of Army Alpha and Army Beta. That several major achievement tests were available before any group intelligence tests were available may be an important point to note in our assessments of the history of educational measurement and educational research. Achievement tests may have prepared the way for group intelligence tests. Educators were committed to testing, and to looking for more efficient ways to test. The possibility of one test rather than a series of test may have been appealing to many simply on the grounds of efficiency.

Another approach to the periodization of the history of American educational research has been offered by the Committee on Educational Research of the National Academy of Education. In its work it is suggested that there are four major periods. The first (1855-1895) is said to begin with Henry Barnard's American Journal of Education. It is

a period characterized as "prescientific" and based on the Wissenschaft ideal. The second period (1895-1938) is seen as having begun with the publication William James' Principles of Psychology and the founding of the Laboratory School at the University of Chicago, and the early works of Rice and Thorndike. It is depicted as "a period of vigorous empiricism, with emphasis on the quantitative measurement of educational effectiveness." The Eight-Year Study, the beginning of the Social Frontier, and the publication of the 1938 NSSE Yearbook on The Scientific Movement in Education mark the beginning of the third period (1938-1955). The fourth period began "when the federal government began to assume a more central role in education, and there was a resurgence of interest in educational problems among scholars outside the ranks of educational specialists."²⁵

The National Academy's approach, like Monroe's, overlooks the strong development of achievement tests before World War I, the development of statistical techniques (factor analysis, for example) and appears to assume that lively debate among educators in the 1930 s and 1940 s about social and economic issues constituted an abandonment of research and measurement. The implicit assumption that those who had been doing quantitative research turned from it to take up social and political theory is a generalization that needs examination. Some, of course,

did do that. Testing did decrease in many public school systems during the 1930 s because money was scarce and mass testing was a budget item that could be easily eliminated without major protest from teachers. However, unless a distinction is made between inquiry on the one hand and popularization of soci-political views and educational criticism on the other, one part of the educational field can be mistaken for the whole. One form of activity may be more dominant than another in a given period, and at later times historians may be more interested in one part than another. However, it is unlikely that one kind of activity gave way completely to the other. We should entertain the possibility that developments occur simultaneously as well as serially.

Charles Brauner offers still another approach to the periodization problem in his discussion of Paul Monroe's Cyclopedia of Education (1911-13). For Brauner the publication of the Cyclopedia marked the beginning of one era and the end of another: "Coming at the close of the long epoch of conceptual giants and at the beginning of the new era of fact-gathering worker-bees, Monroe's Cyclopedia crystallized the most important ideas of educational giants in the purest form at the moment of their richest maturity."²⁶ Brauner's point is not simply that the nineteenth century was an age of generalists and

the twentieth an age of specialists. Rather, he is making a distinction between intellectual styles, or perhaps between changes in what Carl Becker described as "climates of opinion." Becker used this notion to illustrate that frequently it is not the logic of an idea but the prevailing climate of opinion that gives it its force.²⁷ While Brauner's colorful and perhaps perjorative language may be misleading, his point still stands and merits further examination. Changes in intellectual styles, or climates of opinion, are precisely what need to be documented to organize an adequate history of educational measurement and research. In doing so, we should be mindful that our own climate of opinion which frequently prompts us to celebrate the halcyon era of our origins is not allowed to distort our perceptions. Educational measurement and educational research, and more recently educational evaluation, have become fields that are distinct from the history of education. We need to be aware of that distinction if we are to understand what climates of opinion have given ideas force in educational measurement and research.

In 1942 Courtis observed that "there had been five distinct trends or emphases in the history of measurement."²⁸ In his descriptions of the five trends he emphasized the beginnings, did not assign terminal dates to the trends, and offered some assessments of each. His neglect of the

end of each trend may serve as a reminder that the field is one that grew not only in size but also in complexity and that perhaps it is a field that has considered problems, techniques, and concepts simultaneously rather than serially. New developments may draw out attention away from earlier ones but that does not necessarily mean that the earlier ones did not persist. For example, researchers did not abandon achievement tests when group intelligence tests became available. Often they used the intelligence tests as a tool to organize "control" and "experimental" groups to which various achievement tests were then administered to determine the worth of various educational methods.

For Courtis the first phase of educational measurement began with modern psychology, when psychology turned from speculation and toward measurement. This occurred about 1875 when psychologists thought they could measure the newly found individual differences. This phase appeared to dire rather quickly as the psychologists discovered that "the individual differences would not correlate with each other, and one could not use a knowledge of them to predict behavior in new situations."²⁹ However, these developments were brought to the United States and further examined by James McKeen Cattell who developed the forerunner of intelligence tests and trained Thorndike and others to use the new measurement techniques.

Rice's examinations marked the beginning of the second phase. While Rice's contribution was not immediately recognized, his "idea of presenting a given examination under standard conditions to different school systems" was a "forward step in scientific progress and eventually, "the use of controlled conditions of testing" became an important protocol in educational measurement.³⁰ Courtis might have added that Rice, as Scates observed, tried to control variables that educational researchers still try to control: heredity, environment, age, intelligence, and socio-economic status. He might have also added that Rice set the problem that educational researchers, including Courtis himself, addressed for at least two decades. Rice had concluded that the differences in students' performances were due to variations in the quality of teaching. For decades researchers tackled that problem.

The third phase began with the intelligence tests of Theophile Simon and Alfred Binet. This phase was one in which there was a flowering of the study of individual differences, "the fruit of which was the concept of mental age." For Courtis "mental age" was "the greatest single contribution yet made in this field." Of interest to him was the development of several "ages." "We have reading age, height age, physiological age, moral age, etc., and the concept is still being expanded and refined."³¹ The

notion of "ages" was of interest to Courtis because his own work led him to conclude that each individual grew at different rates and that educators had to find a way to determine these rates and then offer education appropriate to the individual's growth rates.

Almost simultaneous with the third phase was the fourth. That was the development of achievement tests. It began with the construction of Stone's Arithmetic Reasoning Test (1908). In the development of that test Stone, according to Courtis, "under Thorndike's guidance, utilized the idea of Rice's controlled examination, but added to it the careful standardization of the questions." Courtis' use of Stone's test, as we have seen, led to the development of his own tests which were designed to allow the establishment of achievement norms for each grade. They were "rate tests composed of material made uniform in terms of objective criteria." The rate test measured skill at a specified level of difficulty but for reasons not then known gave "extremely variable results." However, there was no climate of opinion to sustain rate tests. Thorndike "threw the weight of his influence against rate tests" and "ridiculed them out of court for many persons." Yet, as late as 1942, Courtis still maintained that rate tests represented "an avenue of development largely unexplored" and held that

his efforts "at least resulted in the widespread use of educational tests."³²

Stone's and Courtis' tests were soon followed by the Thorndike Handwriting Scale (1910). It represented the influence of Rice and Cattell who stimulated Thorndike to tackle the problem of calibrating a measuring instrument. Its units were based on the Fullerton-Cattell equal difference theorem. It familiarized "thousands with the idea of units and scales" and prepared the way for the flood of tests that were to appear in the next decade.

In 1912 Leonar P. Ayres offered an alternative form of scaling. It was "the sigma scaling technique which yields variability units." Ayres' technique was quickly accepted, or as Courtis expressed it, "seized popular fancy" and B. R. Buckingham, M. R. Trabue, W. A. McCall, L. L. Thurstone and others soon "extended the technique to include assumed normal distributions, scores in tests, and standardized groups."³³ In the same year M. B. Hillegas' Composition Scale appeared. Soon following were a flood of tests: B. R. Buckingham's Spelling Scale (1913); Thorndike's Achievement in Drawing (1913); Thorndike's Ability in Reading (1914); another Spelling Scale by Ayres (1915); Starch's Reading and Grammar Tests (1915); Starch's Latin Test (1916); the Kansas Silent Reading Test (1916) and Trabue's Completion Test Language Scales (1916).

By 1916 a committee headed by George D. Strayer assembled materials for a NSSE Yearbook, Standards and Tests for the Measurement of the Efficiency of Schools and School Systems, in which attention was given to about a dozen achievement tests as well as to the physiological and physical measurements of children, how tests and scales could be used effectively in the supervision and administration of schools, how to rate school buildings, and chapter by Guy M. Whipple on mental, or intelligence tests. There Whipple indicated that it was then "impossible" to urge "any wholesale adoption of them in public-school systems."³⁴ In 1918 the NSSE devoted another Yearbook to educational measurement, The Measurement of Educational Products. In it there was a bibliography of 345 tests and scales that had been constructed for measuring performance in several areas: Algebra, Arithmetic, Drawing; English Composition, Language, and Vocabulary; English Grammar; French; Geography; Geometry; German; Handwriting; History; Latin; Physics; Reading; Spanish; and Spelling.

Besides the hundreds of achievement tests, there were the newly developed intelligence tests. Individual intelligence, Courtis explained, spread widely after Lewis M. Terman's 1916 publication of the Stanford Revision of the Binet. Of significance is Courtis' observation that "when the United States entered the First World War,

Otis had a group intelligence test about ready for publication, and its use by the Army further popularized intelligence testing."³⁵ His claim that the psychologists had the group intelligence test ready or nearly ready before the war is consistent with a recent conclusion reached by Franz Samelson who has noted that "all through the war activities, there runs a slender thread indicating some conflict about whether the focus of the work should be practical-applied or basic scientific research."³⁷ Before the war both Terman and Robert M. Yerkes had been working on intelligence tests and had sought funds from the General Education Fund to gather "large-scale normative data." One indication that the psychologists used the Army testing program for ends other than the war effort was the nature of the data they collected. Had the tests been solely designed to assist the Army in properly classifying and placing recruits, there may have been no need to ask recruits to provide information about their hometown, state or country of birth. However, they did ask and after the war that data was correlated with the test results.

Some, as we know, used the Army test results to demonstrate that blacks were intellectually inferior to whites, but blacks were not the only ones to be mistreated by the "results" of the Alpha and Beta tests. The tests were also interpreted to show that those who held the more

in either nature or nurture could be expressed as a complex curve. Complex curves, however, could be viewed "as composed of two or more parts, each of which is a portion of some simplex curve."⁴⁷ The laws of growth and the influence of any factor on that growth could all be properly described as "simplex situations" and thus could be expressed with a mathematical formula--the Gompertz formula. Claims that

The Gompertz formula is: $y = kg^{c^g}$ (x stands for the variable time; y stands for the measurement of growth at time x; and g, c, and k are the three constants.

Courtis devised a way to transform this exponential equation into a linear equation: $y_1 = r_1t + s_1$. This transformation relied on the development of an isochron theorem (see below). In the linear equation t stands for units of time and y_1 , s_1 and r_1 are the isochrons for y, g and c in the exponential equation.

conditions were never perfectly constant were not important because "the literature of growth is full of curves whose deviation from mathematically perfect simplex curves is, on average, not more than two or three per cent." That small deviation made his new concept "practical."⁴⁸

One difficulty was that the Gompertz equation was an exponential equation and thus cumbersome, if not difficult,

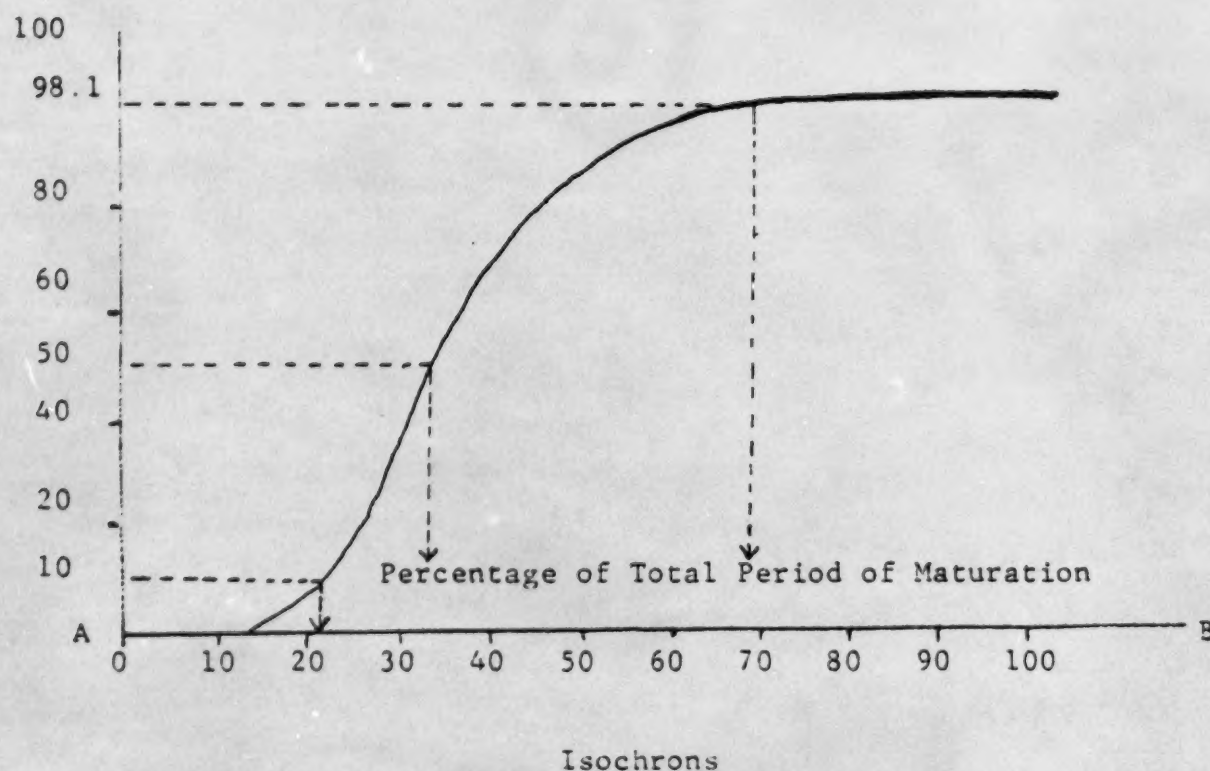
to compute. To overcome this difficulty, he transformed the exponential equation into a linear equation. He was able to do this by observing that the only differences among growth curves, or simplex curves, was the time required for full maturation. Thus, he reasoned that the total time for any and every such growth curve could be divided into one hundred equal units that he called "isochrons." He then claimed it was possible to consider any developments or degrees of growth that were made in equal time intervals to be equal. After applying the isochron theorem to a standard simplex curve (See Figure 1, p. 34), "the development from 0 per cent. to 10 per cent. is considered equal to the development from 10 per cent. to 48 per cent., from 48 per cent. to 80 per cent., etc., since each is made in ten isochrons."⁴⁹

For the remainder of his career and even after his retirement, Courtis continued to write about growth curves and employed his "discovery" in his assessments of the progress and status of educational measurement. He also tried to teach others how to employ the isochron theorem. He prepared a text, The Measurement of Growth (1932) for a course with the same title (C212) that he taught at the University of Michigan. In it he explained the Gompertz law to students, provided a series of exercises to teach

Figure 1

A Mathematically Perfect Curve

Percentage
of
Development



The broken lines indicate the relationship between isochrons and percentage of development.

From: Stuart A. Courtis. *The Measurement of Growth* (Ann Arbor, Michigan: Brumfield and Brumfield, 1932), p. 4.

students how to use isochrons, and even prepared tables of values for isochrons to facilitate computation. Those tables were available for ten cents in stamps. Even in his Philosophy of Education he presented students with discussions of growth, the isochron theorem, and exercises in how to compute with isochrons. There he emphasized that

Scores in no tests can be interpreted correctly except in terms of the individual's own curve of growth. All educational planning must be based on a comprehensive study of the development of the child as a whole because growth in one phase of development may vitally affect or condition growth in other phases.⁵⁰

Courtis' assessment of educational measurement at the end of the 1930 s was that it had just reached the point where it could begin to make significant progress. Its first half century had been characterized by a series of fanciful conclusions based on incorrect assumptions:

It is a fact that we have many instruments of measurement, but our assumptions that they measure intelligence, reading, personality, etc. can easily

be proved to be gross superstitions. It is a fact that we have a number of scaling techniques, but it can readily be shown that these techniques are based on unwarranted assumptions and yield at best only units of performance, not units of ability. It is a fact that our libraries are filled with countless records of experimentation and statistical analyses, having all the form and trappings of scientific procedure, but it is not difficult to show that all this fine appearance is appearance only, that all our experimentation likewise rests on untenable foundations and really does not yield enduring truth.⁵¹

Courtis had not and did not abandon testing and measurement but he did offer many criticism of practices, assumptions, and interpretations of tests. Some of these he directed to his colleagues in measurement and others he directed to educational practitioners (teachers and school administrators) and to parents. Some of his titles in the 1940 s and 1950 s give the impression that he is wants to throw out tests, but he always stopped short of that point.

In a 1938 issue of the Journal of Educational Research Courtis wrote that "after thirty years of study my conclusion

is that no single test and no battery of tests of any type or description yields unambiguous information about the quantities educationalists wished to measure." Current instruments were about as valuable as those of the ancient astrologers and alchemists. Interpretation was "intelligent guess work"--Courtis' term for "scientific deduction."⁵² Progress in interpretation of tests could not be made until two important conditions had been satisfied. First, it was necessary to develop a measure of effort, for it was essential to know how hard an individual was trying in order to interpret his or her test scores. From time to time Courtis brought this factor to the attention of his readers by indicating that sickness may affect performance or by relating how students might be stimulated to work harder or faster by either the promise of a reward or an encouraging talk. Second, Courtis wanted to abandon age and/or grade norms and replace them with repeated measures of the individual's "progress in terms of his own growth curve." ⁵³

As Courtis had outlined at the Indiana Educational Measurements Conference in 1942, the various studies of child growth provided researchers with hundreds of growth curves. All they needed to do was to use them. His own analyses of these studies as well as his own studies convinced him that "the pattern of growth for an individual

differs markedly from the pattern of growth derived from mass averages." Growth was "cyclic in character" and for different individuals the growth occurred at different times.⁵⁴ "Each personality is unique, his scores have a unique interpretation."⁵⁵ Similar or identical scores on a test did not indicate that the children were similar or identical. For example, two children who are equal in height at the same time do not necessarily grow to identical or even similar heights. Similarly, two children who achieve identical scores on a test at the same time may not have the same scores once each has become fully grown or even on subsequent tests. Rates of growth varied and those variations made a difference.⁵⁶

Courtis also objected to the excessive dependence on deductive reasoning in educational research and measurement. Form "the basis of the scientific method is the pragmatic empiricism of the cycle--experimentation, generalization, prediction, and verification." Statistics, as a branch of mathematics, was not an experimental discipline but a deductive one. In the literature on educational measurement, he noted, there was an increasing tendency toward using "logical mathematical reasoning" rather than using "concrete experimental evidence." That tendency was widening the gap between those who constructed the tests and the scales and those who used them--the

teachers. Mathematical reasoning about educational matters needed to be checked by objective experimentation, but it was not being so verified. Moreover, teachers, he warned, "cannot read and understand current discussions of item analysis, factor loadings, correlation pathways, and the host of statistical methods and proofs to be found in measurement articles, but they do sense the contrast between the promises of educational measurement and the actual inadequacies of tests in the classroom."⁵⁷ Courtis did practice what he advocated. His articles are replete with examples taken from tests administered to students or measurements from students taken in school settings. Frequently, they were tests that he had administered himself. However, it would be a mistake to characterize his work as always easy to follow.

Courtis continued to raise questions about how statistics were used in education. He objected to "mass statistics" and called for "personalized statistics." He argued that "research in psychology and education is stymied by the domination of mass statistics in analyses of results of tests and experiments." According to R. A. Fischer, he reminded his readers, any statistical value that was derived from an observed or measured sample was calculated with a specific purpose in mind. That purpose was to make a characterization of the group. The difficulty

was that education was concerned "not with groups, small or large, but with individual children." No two were alike so it made no sense to assume that they were and use "mass statistics."⁵⁸ To do so was to make false assumptions and was not unlike grouping a potato, an apple and an onion, and then performing a series of measurements to determine their characteristics. In fact, Courtis worked such an exercise and reported the values for the mean weight, the variance, and the standard deviation. He explained that he could improve his validity and reliability if he went on to weigh a carload of each and emphasized that increasing the validity and the reliability of the measurements would tell nothing about the significance of the measurements.

Misuse of statistics was not confined to psychology and education. In the late 1930s he related, the Bureau of Home Economics sponsored a WPA Project that was designed to take measures of 36 different body dimensions from nearly 150,000 children in 16 states and the District of Columbia. The directors of the project did make some divisions--sex and age--but then "they averaged tall boys, short boys, fat boys, thin boys, boys of different racial and socio-economic backgrounds, boys in different cycles of physical development, of different levels of health, and of different environmental backgrounds."⁵⁹

That project had essentially averaged potatoes, apples and onions. The project directors recognized that other classifications could have been made but, claimed Courtis, they were too expedient in their approach. To make his point he quoted from the project report: "Although the data from this study (bivariate distributions) indicate that age is a poor basis for sizing garments, the widespread interest in the dimensions of hypothetical average children made it seem important to include in this report an analysis of the distributions of measurements on an age basis."⁶⁰ In other words, there was a climate of opinion that sustained the notion that garments should be marked with an appropriate age and that for each age there was an appropriate set of measurements to be used in manufacturing garments. There was also a climate of opinion that sustained the notion of an "hypothetical average." It cannot be reported whether Courtis wanted every child's measurements taken so that appropriate garments could be assembled, but he did want to change the nature of measurement in education.

Courtis had specific recommendations for how statistics were to be used in educational measurement:

Statistical methods must be used in which averaging correlation, factorial analysis, etc. will not occur until the utmost has been done

to render the original measurements homogeneous by separation into groups as alike as it is possible to make them. Similarly, the actual measurements must be converted into homogeneous groups.⁶¹

To illustrate what he meant he suggested that while it did not make sense to add the weights of the apple, potato and onion, and then compute averages, standard deviations, etc., it did make sense to report that "the potato weighed 2.3 times the onion, and the apple, 1.3 times the onion." Such a procedure would provide comparable units that could be named. They would also provide us with "elements" that would have more meaning than "intelligence, ability, attitudes, etc., upon whose definitions psychologists cannot agree."⁶² As matters stood, the "elements" used by psychologists and educational researchers were comparable to the fire, water, air and earth of the alchemists.

All growth, Courtis firmly believed, could be expressed with the same mathematical formula, but all teachers needed to know was that the rate of growth of an individual child made a difference. The formula itself was "of value to statisticians, not to most teachers." All that teachers needed to know about the growth pattern was that

. . . learning in its early stages proceeds very slowly, but gradually the process speeds up and for a time seems to move forward at a fairly uniform rate. But in the end it inevitably rounds off to a maximum that differs for each individual. Teachers should know, also, that the change from speed-up to slow-down takes place at about a third of the total period of development (marked 'inflection point' on the curve), and that the last stages of mastery are acquired very slowly.⁶³

If teachers knew about the rates of growth and the pattern of growth, there were practical applications they could make to their work. For example, since the typical passing mark--75%--was achieved at the point where an individual had achieved 42% of his or her total growth, there was little advantage in pushing for mastery, say 95%. To achieve mastery it would take longer to gain the additional 20% than it did to achieve the 75% level (58% vs. 42%), for the ability to attain mastery came after the point where growth had slowed somewhat. That, according to Courtis, explained why student performance so frequently fell once drill was curtailed. Rather than invest energy in trying to achieve mastery it would be better to wait and allow nature to take its course.

Teachers, Courtis urged, should study the growth of their children and surrender the notion that they had something to teach. Ideally, teachers' efforts should be devoted "to helping them achieve their purposes, to answering their questions, to rewarding effort sympathetically, to enjoying the dynamic vitality of the children and cooperating with them."⁶³ This notion took Courtis back to the pre-scientific conception of teaching as gardening. In a sentence that reads as though it came from a Jersey Kozinsky novel Courtis instructed teachers: "In the garden where children grow, you, like the farmer, can prepare the soil, water the ground, but you simply cannot 'learn' children at will."⁶⁴ Teachers did not do the growing; children did. It was not even likely that teachers could serve as "guides," for the requirements for effective guidance were not easily attained: "A prerequisite for true guidance is at least several years of developmental measures of growth in many distinct aspects of personality--physical, mental, emotional, social and spiritual as well as educational--and the intimate confidences of the individual."⁶⁵ The best that teachers could do was to stop worshipping tests and scales and opt for the "sympathetic understanding appraisal of the pupil" rather than the "test score" whenever there was a conflict between the two.

Even as late as 1950, Courtis was arguing for a new way to interpret test scores. He then claimed that true comprehension of learning as "primarily a maturation process" worked "havoc with many of the commonly accepted ideas in education."⁶⁶ I.Q. scores, he once again argued, were not indicators of capacity but were merely an indication of the "developmental status of the child." More important than what any test could determine was "some major factor in the maturation process."⁶⁷ Whatever that factor that Courtis could not yet identify was it accounted for the various differences that testing revealed. This unidentified factor was so powerful that "a child's score in a test increases from year to year whether he learns anything or not, just because he is larger, stronger, more mature."⁶⁸

The I.Q. Test and all other tests could be stripped of their "specific factors" and that stripping revealed that one's performance in all tests tended "to approach common values." Because teachers and school administrators generally did not know how to make the detailed analyses necessary for proper interpretation of tests, they were "of little practical value." Clearly, Courtis wanted to curtail the use of tests. He claimed that "most of the conclusions drawn from the results of measurements made with educational and mental tests are either false or

misleading, especially as they relate to evaluation of individuals, teachers or methods of teaching."⁶⁹ The belief that intelligence tests measured intelligence and that spelling tests measured spelling ability was "false" and stood in the way "of an appraisal of factual data from the point of view of maturation."⁷⁰ Courtis had concluded that "the conventional teacher's mind is cluttered with a false ideology" and he was seeking to remove the clutter.

Use of the Gompertz formula and isochrons did not overtake the field of educational measurement. From his point of view it would be fair to state that "mass statistics" still dominates educational measurement and research although educators are probably more sensitive to the limits and possible abuses of testing than they were a few decades ago. Other changes that Courtis advocated also failed to overtake the field of education. After the election of F.D.R. Courtis seemed to believe that social and educational change were possible. Then it was clear that Americans had to reconsider the relationships between freedom and social control. Thus, he reasoned, "it would be wise to give the next generation a reconstructed philosophy of life."⁷¹ His Philosophy of Reconstruction in Public Education called for a form of schooling that would give students practice in democratic skills. The nation needed an alternative to the "prevailing philosophy"

that "has been highly individualistic."⁷² Future education had to be more practical and more political than it had been. Courtis urged that

In the schools of the future more attention must be given to experimentation in practical politics. If young people in training were selecting their own leaders, were working under the leaders they had chosen, and generally were having experiences in the activities of school life, with this form and that form of social control, if their attention were given consciously to the solution of problems raised by the conflicting interests of the individual and society, perhaps the machiner of democratic control might be made to perform successfully those essential functions that are being achieved abroad by tyrannical control.⁷³

Through education people could have the progress made by Stalin and Mussolini without paying the price those dictators exacted. All we needed to do was recognize that "we have been so busy making sure that pupils mastered the knowledge deemed essential for individual competence that we have totally ignored the social skills that enable individuals

to live and work together."⁷⁴ Courtis claimed that there was reason to believe that the mental development of individuals could be better effected "by having children participate in the work of the world than by confining them to academic study of the glories of the past."⁷⁵

American education, as we know, suffered through the 1930 s and the war years without reconstructing its philosophy of education. After the war, there was a call for reform, but it was a charge to return to an emphasis on subject matter. Tradition dies hard. Calls for reform are usually calls to return to a supposedly better era or a call to do better what is being done. Courtis, like other reformers, seems to have failed to recognize that public schools were designed to administer to masses, not to individuals.

Notes

1. Stuart A. Courtis, "Forty Years of Educational Measurement: An Appraisal and a Prophecy," National Elementary Principal (February, 1946), p. 21.

2. Ibid.

3. Arthur B. Moehlman, "Stuart Appleton Courtis: Master Teacher," Nations Schools, Vol. 33, No. 6 (June, 1944), p. 19.

4. Leonard P. "History and Present Status of Educational Measurements" in Guy M. Whipple (ed). Seventeenth Yearbook of the National Society for the Study of Education, Part II: The Measurement of Educational Products (Bloomington, Illinois: Public School Publishing Co., 1918), p. 13.

5. Stuart A. Courtis, "Courtis Tests in Arithmetic: Value to Superintendents and Teachers" in Guy M. Whipple (ed). Fifteenth Yearbook of the National Society for the Study of Education, Part I: Standards and Tests for the Measurement of the Efficiency of Schools and School Systems (Bloomington, Illinois: Public School Publishing Co., 1916), p. 91.

6. Educational Contributions of Stuart A. Curtis:
Listings and Summaries of Published Articles, Books and
Tests through June 1944 (Detroit, Michigan: Detroit Public
Schools, University of Michigan and Wayne University,
1944), p. 45.

For a bibliography of discussions of Curtis' tests
see: B. R. Buckingham, Margaret Doherty and Josephine
MacLachy. Bibliography of Educational and Psychological
Tests and Measurements. U. S. Bureau of Education Bulletin,
No. 55 (Washington, D. C., 1924), pp. 48-52. There, under
the headings of "Curtis standard practice tests in
arithmetic" and "Curtis standard research test, Series B,"
are 104 entries.

7. George S. Counts. Arithmetic Tests and Studies in
the Psychology of Arithmetic. Supplementary Educational
Monographs, Vol. I, No. 4 (Chicago: University of Chicago
Press, 1917), pp. 4-6.

8. Ibid., p. 120.

9. Walter S. Monroe, "Existing Tests and Standards" in Whipple (ed.). Seventeenth Yearbook of the NSSE, Part II, pp. 73-75.

10. Ibid., p. 81.

11. Ibid., p. 89.

12. Ibid., p. 90.

13. Ibid., p. 95.

14. Moehlman, "Stuart Appleton Courtis," p. 19.

15. Courtis, "Forty Years of Educational Measurement," p. 19.

16. Courtis, "Courtis Tests in Arithmetic," p. 91.

17. Stuart A. Courtis, "Research Work in Arithmetic," Educational Administration and Supervision, Vol. III, No. 2 (February, 1917), p. 61.

18. Harold O. Rugg was also among those who developed tests (Handwriting and Algebra) before World War I. He also published Statistical Methods Applied to Education in 1917. In it he discussed those statistical methods that were necessary for interpreting test scores.

19. Stuart A. Courtis, "Contributions of Research to the Individualization of Instruction" in Guy M. Whipple (ed.). Thirty-Seventh Yearbook of the National Society of Education, Part II: The Scientific Movement in Education (Bloomington, Illinois: Public School Publishing Co., 1938), p. 205.

20. Quoted in Stuart A. Courtis, "Fact and Fancy in Educational Measurement," Bulletin of the School of Education, University of Indiana, Vol. XVIII (September, 1942), pp. 12-13.

21. Ibid., p. 13.

22. For a succinct discussion of Rice's work see: Douglas E. Scates, "Fifty Years of Objective Measurement and Research in Education," Journal of Educational Research, Vol. XLI, No. 4 (December, 1947), pp. 241-247. Scates pointed out that Rice "had a number of factors to dispose of before he could come to his point. Naturally the school people would argue that differences in spelling ability were the result of factors over which they had no control. Heredity? Environment? Age? As best one might without intelligence tests and without refined measures of socio-economic status, he analyzed his data with respect to these factors. . . ." (p. 244).

23. After the publication of Hall's "The Contents of the Child's Mind" in the Princeton Review (1883), the questionnaire became a very popular means of collecting data for educational researchers in the United States,

especially for those who worked in the child-study movement. However, it was not a new instrument. The Journal of the Statistical Society of London (October, 1841) carried a report of a study on the condition of education in Bristol that was conducted with the use of a "circular." Henry Barnard had also gathered information about the condition of education in Connecticut with the use of a questionnaire about fifteen years after the Bristol study (American Journal of Education, May, 1856). In Kind und Welt (1856) B. Sigismund reported that he had tried to add to his own observations of children by sending out inquiries to others. The study that served as the model for Hall had been conducted in Berlin in 1879 (F. Bartholomai und Schwabe, "Der Vorstellungskreis der Berliner Kinder beim Eintritt die Schule," Berlin Statistisches Jahrbuch, 1879).

25. Lee J. Cronbach, Patrick Suppes, et. al. (eds). Research for Tomorrow's Schools (New York: The Macmillan Co., 1969), pp. 33-34.

26. Charles J. Brauner. American Educational Theory (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964), pp. 139-140.

27. Carl L. Becker. The Heavenly City of the Eighteenth-Century Philosophers (New Haven, Conn.: Yale University Press, 1932), pp. 4-5.

28. Courtis, "Fact and Fancy," p. 8.

29. Ibid.

30. Ibid.

31. Ibid., p. 9.

32. Ibid.

33. Ibid., p. 10.

34. Guy M. Whipple, "The Use of Mental Tests in the School" in Whipple (ed). Fifteenth Yearbook of the NSSE, Part I, p. 149.

35. Courtis, "Fact and Fancy," p. 10.

36. Franz Samelson, "World War I Intelligence Testing and the Development of Psychology," Journal of the History of the Behavioral Sciences, Vol. 13, No. 3 (July, 1977), p. 277.

37. Robert M. Yerkes (ed.). Psychological Examining in the United States Army (1921). Quoted in Samelson, "World War I Intelligence Testing," p. 279.

38. Whipple, "Use of Mental Tests," pp. 155-156.
39. Ibid., pp. 153-154.
40. Stuart A. Courtis, "What Does the I.Q. Really Measure?" The Nation's Schools, Vol. XI, No. 1 (January, 1933), p. 14.
41. Courtis, "Fact and Fancy," p. 10.
42. Ibid.
43. Stuart A. Courtis, "Teachers Manual for Courtis Standard Practice Tests in Arithmetic" (New York: The World Book Co., 1920), p. 22. Teachers simply needed to count how many students were below standard at a given time and how many of those were at, or above, standard at a subsequent time. Those two numbers and some simple division allowed the determination of teaching efficiency.
44. Courtis, "Forty Years of Educational Measurement," pp. 21-22.
45. Courtis, "Fact and Fancy," pp. 10-11.

46. Ibid., p. 11.

47. Stuart A. Courtis, "Maturation Units for the Measurement of Growth," School and Society, Vol. XXX, No. 777 (November 16, 1929), p. 685.

48. Ibid., p. 686.

49. Ibid., p. 690.

50. Stuart A. Courtis, Philosophy of Education (Ann Arbor, Michigan: Brumfield and Brumfield, 1939), p. 177. Courtis' italics.

51. Courtis, "Fact and Fancy," p. 11.

52. Stuart A. Courtis, "The Interpretation of Scores in Tests and Examinations," Journal of Educational Research, Vol. XXXI, No. 9 (May, 1938), p. 648.

53. Ibid. Courtis' italics.

54. Courtis, "Fact and Fancy," p. 15.

55. Courtis, "Interpretation of Scores," pp. 648-649.

56. Stuart A. Courtis, "The Rate of Growth Makes a Difference," Phi Delta Kappan, Vol. 30 (April, 1949).

57. Stuart A. Courtis, "Past and Present Trends in Educational Measurement," Review of Educational Research, Vol. III, No. 5 (December, 1938), p. 549.

58. Stuart A. Courtis, "Personalized Statistics in Education," School and Society, Vol. LXXX (May, 1955), p. 170.

59. Ibid., p. 171.

60. "Body Measurements of American Boys and Girls for Garment and Pattern Construction," U. S. Department of Agriculture, Misc. Pub. No. 366. Quoted in Courtis, "Personalized Statistics," p. 171.

61. Courtis, "Personalized Statistics," p. 171.

62. Ibid.

63. Courtis, "Rate of Growth Makes a Difference," p. 320. Courtis' Italics.

64. Ibid., p. 321.

65. Stuart A. Courtis, "Let's Stop This Worship of Tests and Scales," The Nation's Schools, Vol. 31, No. 3 (March, 1943), p. 17.

66. Stuart A. Courtis, "Debunking the I.Q.," The Nation's Schools, Vol. 45, (No. 2 (February, 1950), p. 57.

67. Ibid., p. 58.

68. Ibid. Courtis italics.

69. Ibid., p. 59.

70. Ibid., p. 58.

71. Stuart A. Courtis, "A Philosophy of Reconstruction in Public Education," The Nation's Schools, Vol. 13, No. 6 (June, 1934), p. 21.

72. Ibid., pp. 22.

73. Ibid., pp. 22-23.

74. Ibid., p. 22.

75. Ibid., p. 23.